

# 团 体 标 准

T/GDICA 001-2025

## 用于外观专利图像智能检索的 AI 大模型 系统技术规范

(Technical Specifications for AI Large Model Systems Used in Intelligent  
Retrieval of Design Patent Images)

2025 - 12 - 23 发布

2025-12 -24 实施

广东省信息消费协会

发布



## 目 次

前 言 .....	II
1 范围 .....	1
2 规范性引用文件 .....	1
3 术语和定义 .....	1
4 设计原则 .....	2
4.1 高效性 .....	2
4.2 扩展性 .....	2
4.3 准确性 .....	2
4.4 友好性 .....	3
4.5 安全性 .....	3
5 总体设计 .....	3
5.1 架构设计 .....	3
5.2 流程设计 .....	4
6 技术规范 .....	5
6.1 专利数据采集 .....	5
6.2 图像预处理 .....	6
6.3 特征提取 .....	6
6.4 特征向量化 .....	6
6.5 检索匹配 .....	6
6.6 检索结果显示 .....	6
6.7 离线和在线设计 .....	7
7 维护规范 .....	7
7.1 日常维护 .....	7
7.2 系统升级 .....	7
7.3 故障处理 .....	7

## 前 言

本文件按照 GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

本文件由东莞华南设计创新院提出。

本文件由广东省信息消费协会归口管理。

本文件起草单位：东莞华南设计创新院、广东省工信两化融合发展中心、广东工业大学、开普云信息科技股份有限公司、中国科学院自动化所研究所、深圳市恒程创新知识产权代理有限公司、深圳前沿标准技术服务有限公司、沪渝人工智能研究院、广州智联信息咨询有限公司、金鹏电子信息机器有限公司、中通服建设有限公司、东莞广量控股集团有限公司。

本文件主要起草人：王勇、袁宏伟、丁慧、喻玉立、王瑛、杨春宇、朱风华、张宏达、何睿、周金钢、何翠群、李秀芬、边荣国、侯明军。

## 引 言

本次《用于外观专利图像智能检索的AI大模型系统技术规范》团体标准的制定，核心目的在于填补国内外外观专利AI大模型检索领域的标准空白，为该领域建立统一、科学的技术准则；同时立足产业实际需求，为外观专利检索系统的研发、测试、评估与使用提供明确指导依据。

在实际作用层面，该标准不仅能推动按标建设的检索系统提升运行稳定性、检索准确性与技术可靠性，还可有效提高专利检索效率、降低人为检索失误率；从产业发展角度，其落地能促进国内外外观专利AI检索相关技术的迭代进步，进一步为我国技术创新高质量发展提供技术支撑与规范保障，且因制定过程广泛征求了产业上下游、科研院所等多方意见，具备较强的实际应用性与市场推广价值。

该标准在制定过程中广泛征求了产业上下游相关企业、行业协会、研究机构、高等院校的意见，符合实际需求，具有较强的应用性与推广性。



# 用于外观专利图像智能检索的 AI 大模型系统

## 技术规范

### 1 范围

本文件规定了用于外观专利图像智能检索的 AI 大模型系统（以下简称“系统”）的技术规范，包括设计原则、总体架构、模块设计、流程设计、技术规范、测试方法及维护要求。

本文件适用于系统的开发、测试、部署、维护及相关技术评估，可为外观专利检索相关的技术研发、应用实施提供指导。

### 2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 22239—2019 信息安全技术 网络安全等级保护基本要求

### 3 术语和定义

下列术语和定义适用于本文件。

#### 3.1

##### 外观专利

指对产品的形状、图案或者其结合以及色彩与形状、图案的结合所作出的具有美感并适于工业应用的新设计所享有的专利权。

#### 3.2

##### 图像检索

指通过图像处理与识别技术，从图像数据库中找出与查询图像相似或相关的图像的过程。

#### 3.3

##### 外观专利图像检索系统

指专门用于基于图像的外观专利信息检索的计算机系统，包括图像上传、预处理、特征提取、智能检索、结果展示与筛选、专利详情查看与下载以及数据分析与报告生成等功能模块。

#### 3.4

##### AI 大语言模型

基于深度学习技术构建的、具有大规模参数（通常 $\geq 1$  亿参数）和复杂结构的神经网络模型，能够处理

和分析大规模数据，具有强大的特征提取和识别能力。

### 3.5

#### 外观专利图像

指用于表示产品外观设计的专利附图，包含产品的形状、图案、色彩等核心设计特征，是外观专利保护的直观载体。

### 3.6

#### 智能检索系统

利用 AI 技术（如深度学习、自然语言处理等）实现自动化、智能化的信息检索系统，能够根据用户需求（如图像、文本）快速定位相关信息。

### 3.7

#### 特征向量

将图像的关键特征（如形状、纹理、颜色）通过算法转换为的高维数值向量，用于量化描述图像特征，支持相似度计算。

### 3.8

#### 相似度匹配

通过计算查询图像与数据库中图像的特征向量之间的距离（如余弦距离、欧氏距离），判断图像相似程度的过程。

### 3.9

#### 分布式存储

将数据分散存储在多个独立的存储节点上，通过协同机制实现数据的高效存储、访问和扩展的技术架构。

## 4 设计原则

### 4.1 高效性

系统应支持大规模专利图像数据（ $\geq 1000$ 万张）的存储与检索，单条检索请求响应时间 $\leq 2$ 秒；支持 $\geq 100$ 并发用户查询，性能无明显下降。

### 4.2 扩展性

系统架构应采用模块化设计，支持功能模块（如新增特征提取算法、扩展数据源）的灵活增减；支持数据规模的线性扩展，存储容量可随专利数据增长动态扩容。

### 4.3 准确性

检索准确率（正确检索结果占总检索结果的比例） $\geq 90\%$ ，召回率（检索到的相关结果占有所有相关结果的比例） $\geq 85\%$ ；特征提取应能准确捕捉图像的核心设计差异。

#### 4.4 友好性

用户界面应简洁直观，支持图像拖拽上传、关键词联想等便捷操作；检索结果展示应提供多维度筛选（如相似度、申请日期）和可视化分析功能，降低用户使用门槛。

#### 4.5 安全性

符合GB/T 22239—2019中二级及以上安全等级要求，对专利数据、用户信息实行加密存储（采用 AES-256 算法）和访问控制（基于RBAC模型）；防止数据泄露、篡改或非法访问。

### 5 总体设计

#### 5.1 架构设计

系统总体架构分为数据采集层、数据处理层、模型训练层、检索服务层和用户交互层，各层通过标准化接口协同工作，架构如图 1 所示。



图 1 系统总体架构

##### 5.1.1 数据采集层

负责多源外观专利图像及元数据的采集、清洗与标注，具体模块如下：

- 专利图像采集模块：通过 RESTful API 接口对接国家知识产权局、WIPO 等官方数据库，支持批量爬取和增量更新；支持用户上传图像接收与验证；
- 数据清洗模块：去除模糊、重复、无关图像（如分辨率  $< 300 \times 300$  像素的图像），纠正元数据错误（如专利号格式校验）；
- 数据标注模块：为图像添加产品类别（如家电、机械）、设计特征（如曲面、纹理）等标签，标注准确率  $\geq 95\%$ 。

##### 5.1.2 数据处理层

对采集数据进行预处理与特征转化，具体模块如下：

## T/GDICA 001—2025

- 图像预处理模块：对图像进行灰度化、去噪（采用高斯滤波）、增强（对比度调整）、尺寸归一化（统一为 512×512 像素）等处理，预处理后图像清晰度评分≥85 分（采用无参考图像质量评估算法）；
- 特征提取模块：基于预训练的 EfficientNet-B4 模型提取图像深层特征，结合传统算法提取形状（边缘检测）、纹理（LBP 算子）、颜色（HSV 直方图）等浅层特征，形成多模态特征集；
- 特征向量化模块：将多模态特征融合为 512 维特征向量，通过 PCA 降维去除冗余信息，向量存储采用 Milvus 向量数据库。

### 5.1.3 模型训练层

实现 AI 大模型的训练与优化，具体模块如下：

- 模型训练模块：采用≥100 万张标注专利图像数据集进行微调训练，训练过程采用小批量梯度下降（Batch Size=32），迭代次数≥50 轮；
- 模型评估模块：通过准确率、召回率、F1 值等指标评估模型性能，每季度进行一次模型迭代优化，确保性能达标。

### 5.1.4 检索服务层

响应用户检索请求并返回结果，具体模块如下：

- 检索匹配模块：接收用户查询图像，提取特征向量后，采用 ANN 算法结合余弦相似度计算，在特征库中检索相似图像；
- 结果排序模块：综合相似度得分（权重 60%）、专利申请日期（权重 20%）、法律状态（权重 20%）进行排序，优先展示高相关性、近期有效专利；
- 结果展示模块：生成包含图像缩略图、相似度得分、专利号、申请人等信息的结果列表，支持导出为 Excel 或 PDF 格式。

### 5.1.5 用户交互层

提供用户操作入口与反馈机制，具体模块如下：

- 用户注册与登录模块：支持手机号、邮箱注册，采用 OAuth2.0 协议实现第三方登录（如企业微信），登录密码需满足复杂度要求（含大小写字母、数字、特殊字符）；
- 检索请求提交模块：支持图像拖拽上传、拍照上传及关键词辅助检索，提供图像裁剪、旋转等在线编辑工具；
- 结果反馈模块：支持用户对检索结果进行“相关”“不相关”标记，反馈数据用于模型优化；提供意见反馈入口，响应时间≤24 小时。

## 5.2 流程设计

系统检索流程如图2所示，具体步骤如下：

- 数据采集与预处理：从专利数据库采集图像及元数据，经清洗、标注后，通过预处理统一图像格式与质量；
- 特征提取与存储：对预处理图像提取多模态特征，转化为特征向量后存储至向量数据库，建立索引加速检索；
- 用户检索请求：用户上传查询图像，系统对图像进行预处理和特征提取，生成查询特征向量；
- 检索匹配与排序：将查询向量与数据库向量进行相似度计算，按预设规则排序后生成检索结果；
- 结果展示与反馈：以可视化方式展示结果，接收用户反馈并用于系统优化。

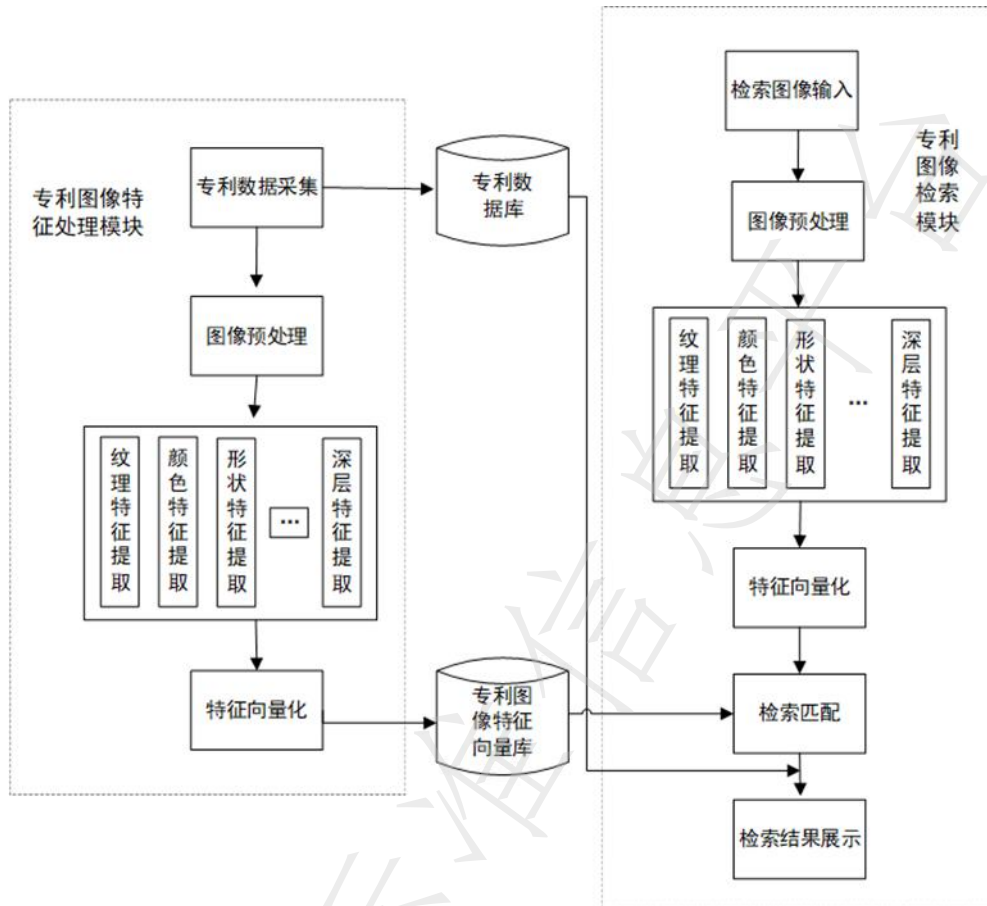


图 2 检索流程

## 6 技术规范

### 6.1 专利数据采集

#### 6.1.1 数据来源

数据来源应满足如下规范：

- 应覆盖国内外主流专利数据库，包括但不限于：国家知识产权局专利数据库、WIPO 全球设计数据库、欧洲专利局（EPO）外观专利库；
- 支持用户上传外观专利图像，上传图像大小 $\leq 20\text{MB}$ ，分辨率 $\geq 300\text{DPI}$ ，格式限 JPEG、PNG、TIFF。

#### 6.1.2 数据采集规范

数据采集规范应满足如下规范：

- 批量采集频率：每日增量采集最新专利数据，每月全量同步一次历史数据，确保数据时效性 $\leq 24$ 小时；
- 数据完整性：采集的元数据应包含专利号、名称、申请人、申请日、授权日、法律状态、产品分类号等核心字段，缺失率 $\leq 1\%$ ；
- 合法性：采集数据需符合《中华人民共和国专利法》《数据安全法》等法律法规，获得合法授权或使用公开数据。

#### 6.1.3 专利数据库设计

专利数据库设计应满足如下规范：

- 采用“关系型数据库+向量数据库”混合架构：MySQL 存储专利元数据，Milvus 存储特征向量，两者通过专利 ID 关联；
- 向量数据库需支持 $\geq 10$  亿级特征向量存储，单条向量查询延迟 $\leq 100\text{ms}$ ，支持动态索引更新。

## 6.2 图像预处理

图像预处理应满足如下规范：

- 去噪处理：采用非局部均值去噪算法，去除图像高斯噪声、椒盐噪声，处理后信噪比（SNR） $\geq 30\text{dB}$ ；
- 增强处理：通过自适应直方图均衡化（CLAHE）提升图像对比度，确保关键设计特征（如边缘、纹理）清晰可辨；
- 归一化处理：统一图像尺寸为  $512 \times 512$  像素，采用双三次插值法缩放，避免特征失真；
- 元数据标注：自动提取并标注专利号、分类号等信息，手动标注准确率 $\geq 98\%$ ，自动标注准确率 $\geq 90\%$ 。

## 6.3 特征提取

特征提取应满足如下规范：

- 深度学习模型：采用 EfficientNet-B4 作为基础模型，在 $\geq 100$  万张外观专利图像数据集上微调，模型参数 $\geq 1.9$  亿；
- 多模态特征融合：融合深层语义特征（来自深度学习模型）与浅层视觉特征（形状、纹理、颜色），提升特征区分度；
- 特征鲁棒性：对图像旋转（ $\pm 30^\circ$ ）、缩放（ $\pm 20\%$ ）、光照变化（ $\pm 30\%$ ）具有适应性，特征匹配准确率下降 $\leq 5\%$ 。

## 6.4 特征向量化

特征向量化应满足如下规范：

- 向量维度：经 PCA 降维后特征向量维度为 512 维，向量元素取值范围  $[0, 1]$ ，数据类型为 float32；
- 向量化稳定性：相同图像多次特征提取生成的向量余弦相似度 $\geq 0.95$ ，确保特征一致性；
- 索引构建：采用 IVF\_FLAT 索引算法构建向量索引，索引构建时间 $\leq 2$  小时/100 万向量，支持增量索引更新。

## 6.5 检索匹配

检索匹配应满足如下规范：

- 相似度算法：采用余弦相似度为主、欧氏距离为辅的混合度量方式，相似度计算耗时 $\leq 50\text{ms}$  /条；
- 检索策略：支持“精确检索 + 模糊检索”双模式，精确检索针对高相似度（ $\geq 90\%$ ）结果，模糊检索覆盖潜在相关结果（相似度 $\geq 60\%$ ）；
- 性能指标：在 1000 万级图像数据集下，单用户检索响应时间 $\leq 2$  秒，并发 100 用户时响应时间 $\leq 5$  秒，准确率 $\geq 90\%$ ，召回率 $\geq 85\%$ 。

## 6.6 检索结果显示

检索结果展示应满足如下规范：

- 展示内容：应包含图像缩略图（ $200 \times 200$  像素）、相似度得分（保留 2 小数）、专利号、名称、申请人、申请日、法律状态等信息；
- 筛选功能：支持按相似度（降序/升序）、申请日（最新/最早）、专利类型（发明/实用新型/外观）、产品分类等维度筛选；

——可视化分析：提供专利分布热力图（按申请地区）、趋势折线图（按申请年份）、分类占比饼图等分析工具，支持导出分析报告。

## 6.7 离线和在线设计

### 6.7.1 离线部分

离线部分应满足如下规范：

- 数据预处理、模型训练、特征向量生成等非实时任务应在离线环境完成，避免占用在线服务资源；
- 离线数据备份：采用“本地备份 + 异地灾备”策略，每日增量备份，每周全量备份，备份数据保存期限 $\geq 1$ 年，恢复时间 $\leq 4$ 小时。

### 6.7.2 在线部分设计

在线部分应满足如下规范：

- 采用微服务架构部署，核心服务（如检索接口、用户交互）支持水平扩展，通过 Kubernetes 实现容器化管理，确保服务可用性 $\geq 99.9\%$ ；
- 负载均衡：采用 Nginx 实现请求分发，支持会话保持，当单节点负载 $\geq 80\%$ 时自动扩容，避免单点故障。

## 7 维护规范

### 7.1 日常维护

日常维护应满足如下规范：

- 每日巡检：检查数据采集任务、服务器负载、数据库状态，生成巡检报告，异常项处理时间 $\leq 2$ 小时；
- 数据维护：每月清理无效数据（如重复图像、过期专利），优化数据库索引，提升查询效率。

### 7.2 系统升级

系统升级应满足如下规范：

- 模型升级：每季度基于新数据微调模型，升级前需通过性能测试，确保准确率提升 $\geq 2\%$ 或召回率提升 $\geq 3\%$ ；
- 功能升级：新增功能需经过单元测试、集成测试和用户验收测试（UAT），测试通过率 $\geq 95\%$ 方可上线。

### 7.3 故障处理

故障处理应满足如下规范：

- 建立三级故障响应机制：一级故障（如系统瘫痪）响应时间 $\leq 30$ 分钟，恢复时间 $\leq 4$ 小时；二级故障（如检索延迟）响应时间 $\leq 1$ 小时，恢复时间 $\leq 2$ 小时；三级故障（如界面异常）响应时间 $\leq 2$ 小时，恢复时间 $\leq 1$ 小时；
- 故障记录：所有故障需记录原因、处理过程、解决方案，形成故障知识库，用于系统优化。